Upper Bound on our Knowledge about Noncommuting Observables for a Qubit System

Yuji Kurotani[1], Takahiro Sagawa[1], and Masahito Ueda[1,2]

[1] *Department of Physics, Tokyo Institute of Technology, Tokyo 152-8551, Japan*
[2] *ERATO Macroscopic Quantum Control Project, JST, Tokyo 113-8656, Japan*

(Dated: February 1, 2008)

A trade-off relation on our knowledge about two noncommuting observables of a qubit system in simultaneous measurement is formulated. The obtained inequality offers a quantitative information-theoretic representation of Bohr's principle of complementarity, and can be interpreted as a trade-off relation on the asymptotic accuracy of the maximum-likelihood estimation of the probability distributions of observables.

Quantum mechanics features two distinct kinds of uncertainty. One is the quantum fluctuations inherent in a measured system, and the other is the noise caused by the process of measurement. Quantum fluctuations prevent us from knowing a quantum system beyond the probability distribution of the measured observable [1], but the probability distribution itself can be accurately determined by means of an appropriate projection measurement. The uncertainty relation between two noncommuting observables such as the position and the momentum originates from those fluctuations [2, 3, 4, 5]. On the other hand, the noise places a limit on the accuracy of simultaneous measurement. It is known, for example, that simultaneous measurements of two noncommuting observables implies that at least one of them cannot be measured without incurring a measurement error [6]. Despite a long history of study [6, 7, 8, 9], however, the fundamental limit to simultaneous measurement of two noncommuting observables has yet to be fully understood.

A classic analysis of the problem was given by Arthurs *et al.* [10, 11] who, under a special condition called unbiasedness, have shown that the lower bound of the uncertainty product of canonically conjugate observables is twice as large as the standard lower bound of $\hbar/2$, where $\hbar$ is the Planck constant divided by 2π. The underlying physics behind this doubling of the lower bound is that, under the condition of unbiasedness, fluctuations of a system's observable and the noise generated in the measurement process become uncorrelated and that they simply add up. More recently, a number of studies on related problems are conducted without invoking the unbiasedness condition and various uncertainty relations are derived [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23].

In this paper, we derive a trade-off relation concerning the measurement accuracy of two noncommuting observables for a qubit system by considering nonideal joint measurements. We also show that our characterization of the measurement accuracy is closely related to the Fisher information [24, 25], which provides the asymptotic accuracy of the maximum-likelihood estimation of the probability distribution of an observable for a finite number of samples. In reality, only a finite number of samples are available [26], which give us only an imperfect informa-

tion about the probability distribution of an observable for an unknown state. The crucial observation made in this paper is that this imperfection is further deteriorated in the case of simultaneous measurement due to noncommutability of the observables.

We first formulate a simultaneous measurement on a qubit system with respect to two observables $\hat{A} = \boldsymbol{n}_A \cdot \hat{\boldsymbol{\sigma}}$ and $\hat{B} = \boldsymbol{n}_B \cdot \hat{\boldsymbol{\sigma}}$, where three-dimensional unit vectors $\boldsymbol{n}_A$ and $\boldsymbol{n}_B$ indicate the directions of the measurements with $\boldsymbol{n}_A \cdot \boldsymbol{n}_B = \cos\theta$ $(0 < \theta < \pi)$, and $\hat{\boldsymbol{\sigma}} \equiv (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ represents of the Pauli matrices. Since both $\hat{A}$ and $\hat{B}$ have eigenvalues of ± 1, they can be represented as $\hat{A} = \hat{P}_A(+) - \hat{P}_A(-)$ and $\hat{B} = \hat{P}_B(+) - \hat{P}_B(-)$, where $\hat{P}_{A(B)}(\pm)$ are projection operators corresponding to observable $\hat{A}$ $(\hat{B})$. We denote as $p_{A(B)}(i)$ the probability distribution of $\hat{A}(\hat{B})$ given by

$$p_\alpha(i) = \text{tr}(\hat{\rho}\hat{P}_\alpha(i)) \ (\alpha = A, B), \tag{1}$$

where $\hat{\rho}$ is the density operator of the measured qubit system. Each simultaneous measurement on the qubit system should yield a pair of outcomes (i, j) $(i, j = \pm)$ for observables $\hat{A}$ and $\hat{B}$, because their eigenvalues are ± 1. We characterize the probability distributions of obtaining (i, j) with a positive operator-valued measure (POVM) $\{\hat{E}(i, j)\}$ [27, 28], where $\sum_{i,j}\hat{E}(i, j) = \hat{I}$, with $\hat{I}$ being the identity operator. The probability distribution $q(i, j)$ of obtaining an outcome (i, j) is given by

$$q(i, j) = \text{tr}(\hat{\rho}\hat{E}(i, j)), \tag{2}$$

and the marginal POVMs are expressed by

$$\hat{E}_A(i) = \sum_{j=+,-} \hat{E}(i, j), \ \hat{E}_B(j) = \sum_{i=+,-} \hat{E}(i, j). \tag{3}$$

The four positive operators $\hat{E}(i, j)$ are, in general, expressed as

$$\hat{E}(i, j) = r_{ij}\hat{I} + \boldsymbol{x}_{ij} \cdot \hat{\boldsymbol{\sigma}}. \tag{4}$$

The requirements that the sum of the four operators equals the identity operator and that all of them be non-

negative are met if and only if

$$\sum_{i,j} r_{ij} = 1, \ \sum_{i,j} \boldsymbol{x}_{ij} = \boldsymbol{0}, \ |\boldsymbol{x}_{ij}| \leq r_{ij}. \tag{5}$$

The marginal POVMs are also expressed as

$$\hat{E}_\alpha(+) = r_\alpha \hat{I} + \boldsymbol{x}_\alpha \cdot \hat{\boldsymbol{\sigma}}, \ \hat{E}_\alpha(-) = \hat{I} - \hat{E}_\alpha(+), \tag{6}$$

where $r_A \equiv r_{++} + r_{+-}$, $r_B \equiv r_{++} + r_{-+}$, $\boldsymbol{x}_A = \boldsymbol{x}_{++} + \boldsymbol{x}_{+-}$, and $\boldsymbol{x}_B = \boldsymbol{x}_{++} + \boldsymbol{x}_{-+}$. The probability distributions of measurement outcomes are given by

$$q_\alpha(i) = \text{tr}(\hat{\rho}\hat{E}_\alpha(i)). \tag{7}$$

In general, $q_A(i)$ and $q_B(i)$ do not respectively coincide with $p_A(i)$ and $p_B(i)$, because the simultaneous measurement entails a measurement error.

In the following, we consider a class of simultaneous measurements called nonideal joint measurement [15, 17]. A simultaneous measurement $\hat{A}$ and $\hat{B}$ belonging to this class satisfies the condition that

$$\hat{E}_\alpha(i) = \sum_{j=+,-} (F_\alpha)_{ij} \hat{P}_\alpha(j), \tag{8}$$

where F_A and F_B are two-dimensional square matrices:

$$F_\alpha = \begin{pmatrix} r_\alpha \pm |\boldsymbol{x}_\alpha| & r_\alpha \mp |\boldsymbol{x}_\alpha| \\ 1 - r_\alpha \mp |\boldsymbol{x}_\alpha| & 1 - r_\alpha \pm |\boldsymbol{x}_\alpha| \end{pmatrix}. \tag{9}$$

It can be shown that $\boldsymbol{x}_A /\!/ \boldsymbol{n}_A$ and $\boldsymbol{x}_B /\!/ \boldsymbol{n}_B$. Moreover, we can show that F_A and F_B are transition-probability matrices, or stochastic matrices, satisfying

$$0 \leq |\det F_\alpha| \leq 1, \ |\det F_\alpha| = 2|\boldsymbol{x}_\alpha|. \tag{10}$$

Note that the nonideality of measurement is characterized only by the transition-probability matrix in the case of a nonideal measurement. A nonideal measurement expressed by Eq. (8) is formally interpreted as a measurement of $\hat{A}(\hat{B})$ to which classical noise, characterized by a transition-probability matrix or a noisy transmission channel, is added. However, in general, the noise in Eq. (8) arises from interactions described by quantum mechanics.

Let us now quantify the accuracy of measurement. Our purpose is to characterize the accuracy in such a manner that it depends only on the process of measurement and not on measured state $\hat{\rho}$. To do this, we focus on the transition-probability matrix F_α, which characterize the noise caused by the measurement process, and define

$$\mathcal{X}_A \equiv (\det F_A)^2, \ \mathcal{X}_B \equiv (\det F_B)^2, \tag{11}$$

where $0 \leq \mathcal{X}_A \leq 1$ and $0 \leq \mathcal{X}_B \leq 1$. We shall refer to $\mathcal{X}_{A(B)}$ as the accuracy of measurement for $\hat{A}$ ($\hat{B}$).

When $\mathcal{X}_A$ or $\mathcal{X}_B$ equals 1, the measurement of $\hat{A}$ or that of $\hat{B}$ can be shown to be a projection measurement. On the other hand, when $\mathcal{X}_A$ or $\mathcal{X}_B$ equals 0, all of the

operators in $\{\hat{E}_A(i)\}$ or $\{\hat{E}_B(i)\}$ are proportional to the identity operator and therefore no information about $\hat{A}$ or $\hat{B}$ can be obtained from the measurement. Note that achieving both $\mathcal{X}_A = 1$ and $\mathcal{X}_B = 1$ is impossible for simultaneous measurement, because a measurement error is unavoidable in at least one of the two noncommuting observables. In fact, we can derive a stronger trade-off relation between $\mathcal{X}_A$ and $\mathcal{X}_B$:

$$\mathcal{X}_A + \mathcal{X}_B - \mathcal{X}_A \mathcal{X}_B \cos^2 \theta \leq 1, \tag{12}$$

where $\theta = \cos^{-1}(\boldsymbol{n}_A \cdot \boldsymbol{n}_B)$. A similar inequality is obtained in Ref. [13, 22] for a special case of $r_A = r_B = 1/2$. The accessible regime for $\mathcal{X}_A$ and $\mathcal{X}_B$ is illustrated in FIG.1 for the case of $\theta = \pi/6$. For example, $\mathcal{X}_A \to 1$ can be achieved only when $\mathcal{X}_B \to 0$, indicating that when we measure one observable without any measurement error, we cannot obtain any information about the other observable. Trade-off relation (12) between $\mathcal{X}_A$ and $\mathcal{X}_B$ implies an upper bound on our knowledge about noncommuting observables for a qubit system.
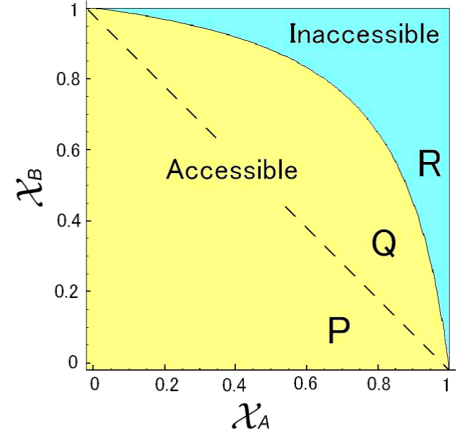

FIG. 1: (Color) Trade-off relation on the accuracy of noncommuting observables. P and Q indicate the regimes satisfying inequality (12) for the case of $\theta = \pi/6$. The regime Q is accessible only through simultaneous measurement.

The derivation of inequality (12) goes as follows. From Eqs. (5), we have

$$\begin{aligned} |\boldsymbol{x}_A + \boldsymbol{x}_B - \boldsymbol{y}| \leq 2r_{++}, \ |\boldsymbol{x}_A - \boldsymbol{x}_B + \boldsymbol{y}| \leq 2r_{+-}, \\ |\boldsymbol{x}_A - \boldsymbol{x}_B - \boldsymbol{y}| \leq 2r_{-+}, \ |\boldsymbol{x}_A + \boldsymbol{x}_B + \boldsymbol{y}| \leq 2r_{--}, \end{aligned} \tag{13}$$

where $\boldsymbol{y} \equiv \boldsymbol{x}_{+-} + \boldsymbol{x}_{-+}$. By using (13) and the triangle inequality, we obtain $|\boldsymbol{x}_A + \boldsymbol{x}_B| + |\boldsymbol{x}_A - \boldsymbol{x}_B| \leq 1$. Noting that $|\det F_A| = 2|\boldsymbol{x}_A|$, $|\det F_B| = 2|\boldsymbol{x}_B|$ and $\boldsymbol{x}_A \cdot \boldsymbol{x}_B = |\boldsymbol{x}_A||\boldsymbol{x}_B|\cos\theta$, we obtain (12).

We give an alternative expression of inequality (12) [22]. Let us define $\mathcal{E}_A \equiv (1/\mathcal{X}_A) - 1$ and $\mathcal{E}_B \equiv (1/\mathcal{X}_B) - 1$, where $0 \leq \mathcal{E}_A < \infty$ and $0 \leq \mathcal{E}_B < \infty$. Here, $\mathcal{E}_A = 0$ implies $\mathcal{X}_A = 1$, that is, the measurement

of observable $\hat{A}$ involves no measurement error. Hence, we can interpret $\mathcal{E}_A$ and $\mathcal{E}_B$ as parameters characterizing measurement errors. In terms of $\mathcal{E}_A$ and $\mathcal{E}_B$, we can express (12) as

$$\mathcal{E}_A \mathcal{E}_B \geq \sin^2 \theta. \tag{14}$$

This inequality manifestly shows that a trade-off relation exists between measurement errors $\mathcal{E}_A$ and $\mathcal{E}_B$ for simultaneous measurement if the corresponding observables are noncommuting (i.e. $\theta \neq 0$).

An optimal simultaneous measurement that satisfies the equality in (12) with $\mathcal{X}_A = \mathcal{X}_B$ is achieved when $\boldsymbol{x}_{++} + \boldsymbol{x}_{--} = 0$, $\boldsymbol{x}_{+-} + \boldsymbol{x}_{-+} = 0$, $|\boldsymbol{x}_{++}| = r_{++} = r_{--}$, $|\boldsymbol{x}_{+-}| = r_{+-} = r_{-+}$, and $r_{++} + r_{+-} = 1/2$. The corresponding POVM $\{\hat{E}(i,j)\}$ is given by $\hat{E}(i,j) = |\boldsymbol{x}_{ij}|\hat{I} + \boldsymbol{x}_{ij} \cdot \hat{\boldsymbol{\sigma}}$ with

$$\begin{aligned} \boldsymbol{x}_{++} = r(\boldsymbol{n}_A + \boldsymbol{n}_B), \ \ \boldsymbol{x}_{-+} = -\boldsymbol{x}_{+-}, \\ \boldsymbol{x}_{+-} = r(\boldsymbol{n}_A - \boldsymbol{n}_B), \ \ \boldsymbol{x}_{++} = -\boldsymbol{x}_{--}, \end{aligned} \tag{15}$$

where $r = \pm \left(|\boldsymbol{n}_A + \boldsymbol{n}_B| + |\boldsymbol{n}_A - \boldsymbol{n}_B| \right)^{-1}/2$. These positive operators are proportional to projection operators, and the directions of the projection are mutually orthogonal [29].

Inequalities (12) and (14) offer a rigorous representation of Bohr's principle of complementarity. According to his classic paper [8], the essence of the principle of complementarity is "the mutual exclusion of any two experimental procedures" when measuring two noncommuting observables simultaneously. Inequalities (12) and (14) quantitatively represent this incompatibility.

We next point out a close connection between the accuracy of measurement $\mathcal{X}_\alpha$ and the asymptotic accuracy of the maximum-likelihood estimation of $p_\alpha(i)$. Let us consider simultaneous measurements for each of N ($< \infty$) samples prepared in the same unknown state $\hat{\rho}$. The issue here is how accurately we can estimate the true probability distributions of $\hat{A}$ and $\hat{B}$ from a finite number of samples. The condition (8) of the nonideal joint measurement is equivalent to the condition that

$$q_\alpha(i) = \sum_{j=+,-} (F_\alpha)_{ij} p_\alpha(j) \tag{16}$$

holds for any $\hat{\rho}$. In this case, probability distributions $q_\alpha(i)$ is parameterized by true probability distributions $p_\alpha(i)$, so that we can estimate $p_\alpha(i)$ by using the method of the classical maximum-likelihood estimation except for the case of $\boldsymbol{x}_A = \boldsymbol{0}$ or $\boldsymbol{x}_B = \boldsymbol{0}$. We assume that $\boldsymbol{x}_A \neq \boldsymbol{0}$ and $\boldsymbol{x}_B \neq \boldsymbol{0}$ in the following analysis.

Suppose that we obtain outcome "+" $N_\alpha(+)$ times and outcome "−" $N_\alpha(-)$ times, where $N_\alpha(+) + N_\alpha(-) = N$. The likelihood function then becomes

$$L_\alpha \left(p_\alpha(+) \right) = \sum_{i=+,-} N_\alpha(i) \ln q_\alpha(i). \tag{17}$$

We denote as p_α^* the maximum-likelihood estimator of $p_\alpha(+)$ inferred from N measurement outcomes such that

$L_\alpha \left(p_\alpha(+) \right)$ takes the maximum value with $p_\alpha(+) = p_\alpha^*$ under condition $0 \leq p_\alpha(+) \leq 1$. The accuracy of the estimate can then be asymptotically characterized by the following theorem [24, 25]: *the distribution of p_α^* approaches the normal distribution with average $p_\alpha(+)$ and variance $(NI_\alpha)^{-1}$ as the number of samples N increases with I_α being the Fisher information.* Note that this theorem does not hold if $p_\alpha(+) = 0$ or 1. In other words, the maximum-likelihood estimator p_α^* asymptotically approaches true value $p_\alpha(+)$, the asymptotic behavior being characterized by Fisher information I_α. The larger the Fisher information, the more information we can extract from the measurement outcome. In the special case of $I_\alpha = 0$, no information about the measured system can be obtained from measurement. Once we have found $p_\alpha(+)$, $p_\alpha(-)$ is obtained from $p_\alpha(-) = 1 - p_\alpha(+)$. In the present situation, the Fisher information is given by

$$\begin{aligned} I_\alpha &\equiv - \sum_{i=+,-} q_\alpha(i) \frac{\partial^2 \ln q_\alpha(i)}{\partial p_\alpha(+)^2} \\ &= \frac{\mathcal{X}_\alpha}{q_\alpha(+)q_\alpha(-)}. \end{aligned} \tag{18}$$

Note that in simultaneous measurement the Fisher information on observable $\hat{A}$ ($\hat{B}$) is deteriorated by a factor of $\mathcal{X}_{A(B)}$ in comparison with projection measurement.

We thus conclude that the trade-off relation (12) or (14) indicates that knowing about one of the probability distributions of the two noncommuting observables prevents us from knowing about the other because the number of available samples is finite; the upper bound on the estimation accuracy of the probability distributions is severely restricted, unless the number of available samples is infinite. Note that if N is infinite, we can accurately reconstruct both $p_A(i)$ and $p_B(i)$ from measured probability distributions $q_A(i)$ and $q_B(i)$; with infinite samples, the noncommutability does not affect the accuracy of the estimates of probability distributions of noncommuting observables.

We have considered the estimates of probability distributions $p_A(i)$ and $p_B(i)$ for simultaneous measurement by using the POVM which is made up by four positive operators. We now discuss a simple strategy for estimating the probability distributions of noncommuting observables. We divide N identically prepared samples into two groups according to the ratio $\xi : 1 - \xi$ ($0 < \xi < 1$), and estimate $p_A(i)$ by measuring $\hat{A}$ for the first group using POVM $\{\hat{E}_A(\pm)\}$, and similarly we estimate $p_B(i)$ for the second group using POVM $\{\hat{E}_B(\pm)\}$. For a large number of samples, this measurement is asymptotically described by POVM $\{\xi\hat{E}_A(\pm), (1-\xi)\hat{E}_B(\pm)\}$. In this case, the accuracy of the measurement is limited by $\mathcal{X}_A \leq \xi$ and $\mathcal{X}_B \leq 1 - \xi$, because the accuracy per sample is deteriorated by a factor of ξ and $1 - \xi$, respectively. We thus obtain $\mathcal{X}_A + \mathcal{X}_B \leq 1$ (domain P in FIG.1), and therefore conclude that simultaneous measurement has an advantage over this simple method in that the domain $\mathcal{X}_A + \mathcal{X}_B > 1$ for $\theta \neq \pi/2$, i.e. domain Q in FIG.1

is accessible.

It is worth pointing out that we can interpret the trade-off relations (12) and (14) as uncertainty relations between the measurement error and the back-action of the measurement [7, 16, 21, 30, 31, 32, 33]. Let us suppose that $\hat{\rho}'$ is a state immediately after measurement of $\hat{A}$ for premeasurement state $\hat{\rho}$. To identify the disturbance of $\hat{B}$ caused by $\hat{A}$, we consider how much information about $\hat{B}$ for premeasurement state ρ is left in postmeasurement state $\hat{\rho}'$. We characterize this by considering how much information on $\hat{\rho}$ can be obtained by measuring $\hat{B}$ for $\hat{\rho}'$. The joint operation of measurement $\hat{A}$ being followed by measurement $\hat{B}$ can be described by a POVM $\{\hat{E}(i,j)\}$ corresponding to the measurement outcomes (i,j) [28]. If the POVM fulfills the requirement that its marginal POVMs describe nonideal measurements of $\hat{A}$ and $\hat{B}$, we again obtain inequality (12). It is possible to interpret $1 - \mathcal{X}_B$ as a measure of the back-action of $\hat{B}$ caused by measurement of $\hat{A}$. Defining the measurement error of $\hat{A}$ as $\mathcal{E}_A \equiv (1/\mathcal{X}_A) - 1$ and the back-action of the measurement on $\hat{B}$ as $\mathcal{D}_B \equiv (1/\mathcal{X}_B) - 1$, we obtain

$$\mathcal{E}_A \mathcal{D}_B \geq \sin^2 \theta. \tag{19}$$

We should note that a non-selective measurement process for $\hat{A}$ can simulate the decoherence caused by the environment. In this case, trade-off relations (12) and (19) give a lower bound to the back-action of $\hat{B}$ in the presence of decoherence characterized by $\mathcal{X}_A$.

Note that if it were possible to perform error-free measurement of $\hat{A}$ without disturbing $\hat{B}$, then we could precisely measure both $\hat{A}$ and $\hat{B}$ by performing measurement $\hat{A}$ followed by measurement $\hat{B}$. We can say that the formal similarity of two kinds of trade-off relations, (12) (or equivalently (14)) and (19), quantitatively represents the above-mentioned connection between the two kinds of uncertainty relations [16, 20]: the uncertainty relation between the error of $\hat{A}$ and that of $\hat{B}$, and the uncertainty relation between the error of $\hat{A}$ and the back-action of $\hat{B}$. In the context of the maximum-likelihood estimation with finite samples, the number of samples required to achieve a given accuracy in estimating the probability distributions of both $\hat{A}$ and $\hat{B}$ is greater than the number for the case of estimating either $\hat{A}$ or $\hat{B}$, because samples on which we perform precise measurements of $\hat{A}$ lose the information about the probability distribution of $\hat{B}$, due to the back-action of measurement $\hat{A}$.

Finally, we mention the relevance of our work to previous work in the case of a qubit system. Arthurs and Goodman [11] have discussed simultaneous measurement by assuming the unbiasedness for observables $\hat{A}$ and $\hat{B}$. This condition implies that the arithmetic average of the measurement outcomes approaches to the true average in the limit of $N \to \infty$. In the case of a qubit system, the unbiasedness condition implies $\mathcal{X}_A = \mathcal{X}_B = 1$ because the probability distribution has a one-to-one correspondence with the average value. As discussed earlier, however, no simultaneous measurement satisfies this condition. We therefore cannot address simultaneous measurement by the method of Arthurs and Goodman for a qubit system. On the other hand, Andersson et al. [22] have relaxed the unbiasedness condition and adopted the condition that the arithmetic average of the measurement outcomes is proportional to the true average in the limit of $N \to \infty$. This condition is satisfied if and only if $r_A = r_B = 1/2$. In this paper, we have considered the most general class of measurements to which we can use the maximum-likelihood estimation to the marginal probability distributions, without assuming the unbiasedness condition. This class of measurements has been studied by de Muynck and Martens who have derived a trade-off relation in terms of the Shannon channel capacity [15, 17].

In conclusion, we have derived a trade-off relation concerning the accuracy of simultaneous measurement of two noncommuting observables for a qubit system. The relation gives an upper bound on our knowledge about noncommuting observables. Moreover, we have pointed out that the accuracy parameter $\mathcal{X}_{A(B)}$ is quantitatively related to the asymptotic accuracy of the maximum-likelihood estimation of the probability distribution of observable $\hat{A}(\hat{B})$. The generalization of our results to high-spin systems and continuous-variable systems merits further study.

---

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] E. H. Kennard, Z. Phys. **44**, 326 (1927).
[3] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[4] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
[5] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
[6] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton, 1955).
[7] W. Heisenberg, Z. Phys. **44**, 172 (1927).
[8] N. Bohr, Phys. Rev. **48**, 696 (1935).
[9] M. Jammer, *The philosophy of quantum mechanics* (Wiley, New York, 1974).
[10] E. Arthurs and J. L. Kelly, Jr., Bell. Syst. Tech. J. **44**, 725 (1965).
[11] E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).
[12] H. P. Yuen and M. Lax, IEEE Trans. Inf. Theory **IT-19**, 740 (1973).
[13] P. Busch, Phys. Rev. D **33**, 2253 (1986).
[14] Y. Yamamoto and H. A. Haus, Rev. Mod. Phys. **58**, 1001

(1986).

[15] H. Martens and W. M. de Muynck, Found. Phys. **20**, 255 (1990).

[16] D. M. Appleby, Int. J. Theor. Phys. **37** 2557 (1998).

[17] W. M. de Muynck, Found. Phys. **30**, 205 (2000).

[18] P. Busch and C. R. Shilladay, Phys. Rev. A **68**, 034102 (2003).

[19] M. J. W. Hall, Phys. Rev. A, **69**, 052113 (2004).

[20] M. Ozawa, Phys. Lett. A **320**, 367 (2004).

[21] R. Werner, Qu. Inf. Comp. **4** 546 (2004).

[22] E. Andersson, S. M. Barnett and A. Aspect, Phys. Rev. A **72**, 042104 (2005).

[23] S. Massar, e-Print: quant-ph/0703036 (2007).

[24] R. A. Fisher, Proc. Camb. Phil. Soc. **22**, 700 (1925).

[25] T. M. Cover and J. A. Thomas, *Elements of Information theory* (John Wiley and Sons, 1991).

[26] S. Massar and S. Popescu, Phys. Rev. Lett. **74** 1259 (1995).

[27] E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).

[28] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[29] L. Cohen and M. O. Scully, Found. Phys. **16**, 295 (1986).

[30] C. A. Fuchs, Fortschr. Phys. **46**, 535 (1998).

[31] K. Banaszek and I. Devetak, Phys. Rev. A **64**, 052307 (2001).

[32] M. Ozawa, Ann. Phys. **311**, 350 (2004).

[33] P. Busch, T. Heinomen and P. J. Lahti, Phys. Lett. A **320**, 261 (2004).